

CLAIMS

1. A method of searching data in databases using an ensemble of models, said method comprising:
 - ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;
 - selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy; and
 - applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.
2. The method in claim 1, wherein said sub-ensemble includes fewer models than said ensemble.
3. The method in claim 1, wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble.
4. The method in claim 1, wherein, the size of each sub-ensemble is different and has a potentially different level of confidence.
5. The method in claim 1, wherein the size of said ensemble is fixed.
6. The method in claim 1, wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process.
7. The method in claim 1, further comprising, before said selecting, calculating confidence values of different sub-ensembles.

8. A method of searching data in databases using an ensemble of models, said method comprising:

ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy, such that said sub-ensemble include only the most accurate models; and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.

9. The method in claim 8, wherein said sub-ensemble includes fewer models than said ensemble.

10. The method in claim 8, wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble.

11. The method in claim 8, wherein the size of each sub-ensemble is different and has a potentially different level of confidence.

12. The method in claim 8, wherein the size of said ensemble is fixed.

13. The method in claim 8, wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process.

14. The method in claim 8, further comprising, before said selecting, calculating confidence values of different sub-ensembles.

15. A method of searching data in databases using an ensemble of models, said method comprising:

performing training comprising:

ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

joining different numbers of models together to form sub-ensembles, wherein models are joined together in said sub-ensemble in said order of prediction accuracy;

calculating confidence values of each of said sub-ensembles; and

making a prediction comprising:

selecting a sub-ensemble of said models that meets a given level of confidence; and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.

16. The method in claim 15, wherein said sub-ensemble includes fewer models than said ensemble.

17. The method in claim 15, wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble.

18. The method in claim 15, wherein the size of each sub-ensemble is different and has a potentially different level of confidence.

19. The method in claim 15, wherein the size of said ensemble is fixed.

20. The method in claim 15, wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process.

21. A service of searching data in databases using an ensemble of models, said service comprising:

ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy; and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.

22. The service in claim 21, wherein said sub-ensemble includes fewer models than said ensemble.

23. The service in claim 21, wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble.

24. The service in claim 21, wherein the size of each sub-ensemble is different and has a potentially different level of confidence.

25. The service in claim 21, wherein the size of said ensemble is fixed.

26. The service in claim 21, wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process.

27. The service in claim 21, further comprising, before said selecting, calculating confidence values of different sub-ensembles.

28. A program storage device readable by a computer tangibly embodying a program of instructions executable by said computer for performing a method of searching data in databases using an ensemble of models, said method comprising:

ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy; and

applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.

29. The program storage device in claim 28, wherein said sub-ensemble includes fewer models than said ensemble.

30. The program storage device in claim 28, wherein said confidence is a measure of how closely results from said sub-ensemble will match results from said ensemble.

31. The program storage device in claim 28, wherein the size of each sub-ensemble is different and has a potentially different level of confidence.

32. The program storage device in claim 28, wherein the size of said ensemble is fixed.

33. The program storage device in claim 28, wherein as the level of confidence is raised, a sub-ensemble that has more models will be selected in said selecting process, and as the level of confidence is lowered, a sub-ensemble that has fewer models will be selected in said selecting process.

34. The program storage device in claim 28, further comprising, before said selecting, calculating confidence values of different sub-ensembles.

35. A system for searching data in databases using an ensemble of models, said method comprising:

means for ordering models within said ensemble in order of prediction accuracy, with the most accurate model being first in said order;

means for selecting a sub-ensemble of said models that meets a given level of confidence, wherein models are joined together in said sub-ensemble in said order of prediction accuracy; and

means for applying said sub-ensemble, in place of said ensemble, to an example to make a prediction.